

## Data Clearinghouse vs. Data Sharing Cooperative

*The Indiana GIS Initiative Data Sharing Committee is investigating models for data sharing that may be desirable for modification and adoption in Indiana. This document represents an overview comparison of a Data Clearinghouse vs. Data Sharing Cooperative. This document was created by the Data Sharing Committee for discussion at the October 19, 2000 Indiana GIS Initiative Quarterly Meeting.*

What is Clearinghouse? (referred to as federal model on

[http://www.nysgis.state.ny.us/coop\\_add.htm](http://www.nysgis.state.ny.us/coop_add.htm))

Fact sheet describing Clearinghouse at

<http://www.fgdc.gov/publications/documents/clearinghouse/chouse.pdf>

The Federal Geographic Data Committee definition of Clearinghouse is a distributed, electronically connected network of geospatial data producers, managers and users. It is a means for locating geospatial data, determining its relevance, and obtaining or ordering it as economically as possible. One of the essential requirements of the Clearinghouse is to support searches for geospatial data over the Internet. The Clearinghouse is not a data archive. The NYS Cooperative website indicates that this federal model requires participants to release their data into the public domain. Also, the federal model has no formal coordination of maintenance or custodial responsibilities for a specific dataset. Any user can distribute the data to another user without restrictions.

### ***Freedom of Information Considerations Concerning a Clearinghouse Direction for Data Sharing versus a Cooperative Direction***

Indiana Access to Public Records (APR) Law seemingly does not prohibit the creation of a GIS Data Sharing Cooperative modeled after New York State. A **distributed** cooperative model would still allow for any person\* to be able to obtain GIS datasets from any public agency since the cooperative would not “own” the datasets. The datasets would still be in the custody of the public agency that developed them. Any said agency would still be required to follow appropriate sections of the APR law in responding to any request to inspect the dataset or obtain the dataset. But this would be beyond the responsibilities of the cooperative.

The NYS model requires that all state agencies participate in the cooperative but only encourages local entities to join. If this were to be the case in Indiana, some form of legislation or regulation would have to be enacted. Some of the sections of the APR law would dictate how the cooperative would operate, e.g. fee issues and how to regulate situations such as if an agency is using another agency’s dataset and a 3<sup>rd</sup> agency wants to use them. Other issues such as licensing and/or copyright of datasets would have to be worked out.

The creation of a distributed clearinghouse would be allowed for the same reasons and with the same restrictions.

The general conclusion is that the Data Sharing Committee could move forward with either model, but the cooperative would be much more complex both practically and legally.

---

\*(From IC 5-14-3-2) "Person" means an individual, a corporation, a limited liability company, a partnership, an unincorporated association, or a governmental entity.

### **Applicable sections of the Indiana Code vis-à-vis questions concerning open access for GIS Data Sharing.**

*Note: Any text below that is in **bold** is not emboldened within in the actual code, but is done so by me to highlight certain aspects of the code.*

#### **IC 5-14-3-1**

Sec. 1. A fundamental philosophy of the American constitutional form of representative government is that government is the servant of the people and not their master. Accordingly, it is the public policy of the state that all persons are entitled to full and complete information regarding the affairs of government and the official acts of those who represent them as public officials and employees. Providing persons with the information is an essential function of a representative government and an integral part of the routine duties of public officials and employees, whose duty it is to provide the information. This chapter shall be liberally construed to implement this policy and place the burden of proof for the nondisclosure of a public record on the public agency that would deny access to the record and not on the person seeking to inspect and copy the record.

*As added by P.L.19-1983, SEC.6. Amended by P.L.77-1995, SEC.1.*

#### **IC 5-14-3-3**

Sec. 3. (a) Any person may inspect and copy the public records of any public agency during the regular business hours of the agency, except as provided in section 4 of this chapter. A request for inspection or copying must:

- (1) identify with reasonable particularity the record being requested; and
- (2) be, at the discretion of the agency, in writing on or in a form provided by the agency.

No request may be denied because the person making the request refuses to state the purpose of the request, unless such condition is required by other applicable statute.

(b) A public agency may not deny or interfere with the exercise of the right stated in subsection(a). The public agency shall either:

- (1) provide the requested copies to the person making the request; or
- (2) allow the person to make copies:

(A)on the agency's equipment; or

(B)on his own equipment.

(c) Notwithstanding subsections (a) and (b), a public agency may or may not do the following:

(1) In accordance with a contract described in section 3.5 of this chapter, permit a person to inspect and copy through the use of enhanced access public records containing information owned by or entrusted to the public agency.

(2) Permit a governmental entity to use an electronic device to inspect and copy public records containing information owned by or entrusted to the public agency.

(d) Except as provided in subsection (e), a public agency that maintains or contracts for the maintenance of public records in an electronic data storage system shall make reasonable efforts to provide to a person making a request a copy of all disclosable data contained in the records on paper, disk, tape, drum, or any other method of electronic retrieval if the medium requested is compatible with the agency's data storage system.

**This subsection does not apply to an electronic map\* as defined by IC 5-14-3-2).**

(e) A state agency may adopt a rule under IC 4-22-2, and a political subdivision may enact an ordinance, prescribing the conditions under which a person who receives information on disk or tape under subsection (d) may or may not use the information for commercial purposes, including to sell, advertise, or solicit the purchase of merchandise, goods, or services, or sell, loan, give away, or otherwise deliver the information obtained by the request to any other person (as defined in IC 5-14-3-2) for these purposes. **Use of information received under subsection (d) in connection with the preparation or publication of news, for nonprofit activities, or for academic research is not prohibited.** A person who uses information in a manner contrary to a rule or ordinance adopted under this subsection may be prohibited by the state agency or political subdivision from obtaining a copy or any further data under subsection (d).

### **IC5-14-3-8.5**

Sec. 8.5. (a) The fiscal body of a political subdivision having a public agency that charges a fee under section 8(j) of this chapter shall adopt an ordinance establishing an electronic map generation fund. The ordinance must specify that the fund consists of fees collected under section 8(j) of this chapter. The fund shall be administered by the public agency that collects the fees.

(b) The electronic map generation fund is a dedicated fund with the following purposes:

(1) The maintenance, upgrading, and enhancement of the electronic map.

(2) The reimbursement of expenses incurred by a public agency in supplying an electronic map in the form requested by a purchaser.

*As added by P.L.58-1993, SEC.9.*

(j) Except as provided in subsection (k), a public agency may charge a fee, uniform to all purchasers, for providing an electronic map that is

based upon a reasonable percentage of the agency's direct cost of maintaining, upgrading, and enhancing the electronic map and for the direct cost\*\* of supplying the electronic map in the form requested by the purchaser. If the public agency is within a political

subdivision having a fiscal body, the fee is subject to the approval of the fiscal body of the political subdivision.

(k) The fee charged by a public agency under subsection (j) to cover costs for maintaining, upgrading, and enhancing an electronic map shall be waived by the public agency if the electronic map for which the fee is charged will be used for a noncommercial purpose, including the following:

- (1) Public agency program support.
- (2) Nonprofit activities.
- (3) Journalism.
- (4) Academic research.

*As added by P.L.19-1983, SEC.6. Amended by P.L.54-1985, SEC.6; P.L.51-1986, SEC.2; P.L.58-1993, SEC.7; P.L.78-1995, SEC.1; P.L.151-1999, SEC.1.*

\*(as defined in IC 5-14-3-2) "electronic map" means copyrighted data provided by a public agency from an electronic geographic information system.

\*\* (as defined in IC 5-14-3-2) "Direct cost" means one hundred five percent (105%) of the sum of the cost of:

- (1) the initial development of a program, if any;
- (2) the labor required to retrieve electronically stored data; and
- (3) any medium used for electronic output;

for providing a duplicate of electronically stored data onto a disk, tape, drum, or other medium of electronic data retrieval under section 8(g) of this chapter, or for reprogramming a computer system under section 6(c) of this chapter.

### ***Broad Participation***

Whatever choice for dissemination of data that the Indiana Geographic Information Council and Indiana GIS Initiative prefer and ultimately choose for implementation, broad participation is an over-arching goal. The INGISI Website should provide the portal to GIS data in Indiana. The question that should be asked of each option is: What is the incentive to participate?

### ***Shared Maintenance***

Shared maintenance in a data cooperative implies a responsibility on both the custodian of the dataset, and the party using it. Cooperative agreements could require that any party who discovers an error in a dataset, or makes enhancements that would be beneficial to all users, reports those changes back to the custodian for inclusion in future releases of the dataset.

There are fewer, or no formal agreements in a data clearinghouse. Data is posted for use with an implied 'caveat emptor', or 'buyer beware'. There may be a mechanism for contacting the data custodian such as a phone number, mailing address, or email address in the metadata, if such exists. Data users can freely copy the available datasets without any obligation to report errors, or inform the custodian that they are using the dataset.

## ***Simplified Sharing***

Simplified sharing in a cooperative allows for the data to be shared at little or no cost to the participants, which aids in minimizing the duplication of effort in creating datasets that are maintained by other agencies. All the members of the cooperative sign the same agreement only once and are allowed to voluntarily terminate their membership upon 90 days written notice to the other members. Upon signing, the member does not have to contribute data, but is allowed access to the other members datasets. As part of the agreement any changes to the data used shall be posted back to the owner of the original data set. This method does not use a data warehouse, but directs the users to the site which contains the data they wish to obtain.

There are some problems that may arise from this type of sharing. The quality of the data could be of questionable nature, unless standards are a part of the submission process. This may lead to a fee structure in order to maintain a standardization and policing committee of the cooperative.

## ***Fees: NYS Cooperative model vs. Federal Clearinghouse model***

The Federal model openly distributes data at cost of duplication to the public with no restrictions. The Cooperative model allows the primary dataset custodian to openly distribute data among members at cost of duplication or less, but charge commercial users “fair market value”. This gives the primary custodian leverage to negotiate with entities outside the Cooperative for value updates to their datasets. Non-members of the Cooperative are assured of better quality data as they receive the geodata directly from the primary custodian, not an intermediary that may have modified the data. For example, the Department of Transportation may distribute its street dataset to a business at no cost provided the business updates the street dataset and gives DOT a copy of the improved data. Neither the Federal or Cooperative models are revenue-generating businesses. The Cooperative does not charge membership fees nor require members to share datasets.

## ***Motivation to Contribute***

Indiana GIS users are questioning the roles and responsibilities of data use and ownership in the given age of Internet data sharing. A significant down-side of the federal model is that there is no formal coordination of maintenance or custodial responsibilities for a given GIS dataset. The open aspect of the model encourages proliferation of datasets to many different users at minimum cost, but the user has no assurance of receiving the most up-to-date version, or more importantly, of receiving the corrections or updates made to the dataset by other users. This is particularly true of datasets downloadable from the Internet. The result is that many users will independently correct or update the dataset to meet their needs. This duplication of effort is very costly and wasteful. One need only consider the condition of the US Census Bureau's TIGER file and wonder how many different users, in both the public and private sectors, have spent time, energy, and money making independent and uncoordinated updates. TIGER data

may be 'free', but what is the real cost of making it useable, and what has been spent, collectively, by everyone independently making it useable?

The Cooperative can have broader participation because it uses a blend of the open sharing aspects of the federal model and licensing provisions of the marketplace models. Open sharing occurs within the Cooperative, with costs never exceeding duplication fees. This follows the federal model. An improvement, however, is that data requests are always referred to the Primary Custodian of the dataset, so the requestor can be confident that they are receiving the most up-to-date and corrected version. A standard NYS GIS Cooperative Data Sharing Agreement (license) is executed once by each member of the Cooperative, thus ensuring uniform 'rules' of data sharing and data maintenance for Cooperative members. The concept of Primary and Secondary Custodians ensures that clear and binding lines of responsibility exist for shared maintenance, revision, and redistribution of the data.

Shared maintenance is perhaps the most significant and cost-saving benefit of the cooperative concept. The shared maintenance is not even a burden for Cooperative members; the Cooperative only requires that any error corrections or revisions that are performed on the data be shared back with the Primary Custodian. This mechanism harnesses the data maintenance actions that would happen anyway, to produce better data at lower composite cost for all Cooperative members. Primary Custodians will need to evaluate the contributed revisions and manage the overall updates of their datasets. In many cases, the contributed changes and revisions could serve as "triggers" to indicate areas in the database requiring a blanket update, such as from orthoimagery.

Guidelines for participation in the NY Data Sharing Cooperative may be considered controversial – membership is limited to non-commercial entities. This policy should be reviewed by the Indiana GIS community for acceptability and legality. With the NY model, partnerships with entities outside of the Cooperative are fostered. The ability of Cooperative members to license their data means that, unlike the federal model, members will be able to bargain in good faith as equals with outsiders for the return of value-added data to Cooperative members. Outsiders, either in the private sector or other non-members of the Cooperative, will have the confidence of receiving higher quality data from a Primary Custodian, while Cooperative members will be able to negotiate for value in return for the data, rather than a give-away. The fact that the data is not released into the public domain also removes a barrier that has prevented some private sector entities from partnering with the public sector on joint data development.

### ***Empowered Custodians***

The Cooperative relies on empowered custodians to work. There is no hierarchical organization to the Cooperative, but rather a network of data custodians operating as equals. The terms and conditions of the Data Sharing Agreement reinforce the custodial responsibilities vested in Primary Custodians. Key characteristics of Primary Custodians include:

- Autonomy to decide how and whether or not to incorporate improved data received from Secondary Custodians. An agency is Primary Custodian of a dataset by virtue of having created it, and as such is presumed to have a business need to keep it maintained. They can use their own judgment on determining the best way to do this.
- Sole source for obtaining a particular dataset. Since each dataset will have a single Primary Custodian, and since Secondary Custodians are not permitted to redistribute data obtained from Primary Custodians, there will be no ambiguity on where to obtain a dataset. Likewise, there will be no ambiguity as to whether the most recent version is being obtained.
- Ability to opt for public domain distribution. The Cooperative does not require Primary Custodians to limit distribution of their datasets to the Cooperative. Members who desire to release their own data into the public domain are free to do so. In doing so, however, they will not be able to take advantage of many of the benefits of the Cooperative, since data in the public domain is available to all with no restrictions or obligations.
- Bargaining leverage to negotiate with outsiders for value-added improvements to the data. This enables members to obtain real value in exchange for their data, potentially lowering their costs for improvements that they might otherwise need to purchase or do without.

In summary, the Cooperative does not prescribe to Primary Custodians what they must do with their own data (other than making it available to other Coop Members and adhering to standards), but rather tells them what obligations and limitations they have with other members' data.